

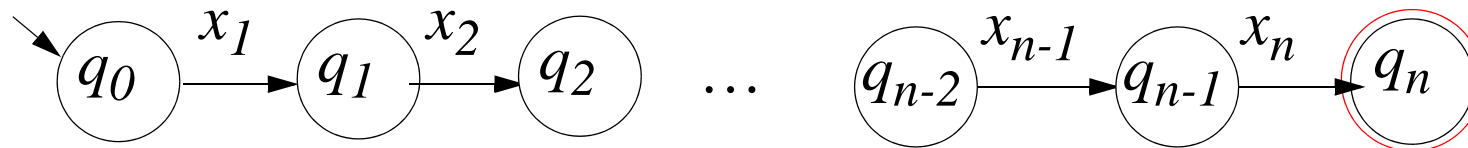
5-B Finite Automata and Regular Grammars

Consider a finite automaton $A = (Q, T, \delta, q_0, F)$ where $\delta: Q \times T^* \rightarrow 2^Q$.

Assume $1 \leq \forall i \leq n: q_i \in \delta(q_{i-1}, x_i), q_i \in Q, x_i \in T^*, q_n \in F$. Then

$q_1 \in \delta(q_0, x_1), q_2 \in \delta(q_1, x_2), \dots, q_n \in \delta(q_{n-1}, x_n), q_n \in F$ or

$q_n \in \delta(\delta(\dots\delta(\delta(q_0, x_1), x_2), \dots, x_{n-1}), x_n) = \delta^n(q_0, x_1x_2\dots x_n) \in F$.



Consider a grammar $G = (N, T, P, S)$.

Assume $1 \leq \forall i \leq n: A_{i-1} \rightarrow x_i A_i \in P, A_i \in N, x_i \in T^*, A_n \rightarrow \epsilon \in P$. Then

$A_0 \Rightarrow x_1 A_1 \Rightarrow \dots \Rightarrow x_1 x_2 \dots x_{n-1} A_{n-1} \Rightarrow x_1 x_2 \dots x_{n-1} x_n A_n \Rightarrow x_1 x_2 \dots x_{n-1} x_n$.

$A_0 \Rightarrow^n x_1 x_2 \dots x_{n-1} x_n A_n \Rightarrow^{A_n \rightarrow \epsilon} x_1 x_2 \dots x_{n-1} x_n = x \in T^*$.

Consider a grammar $G = (N, T, P, S)$ where

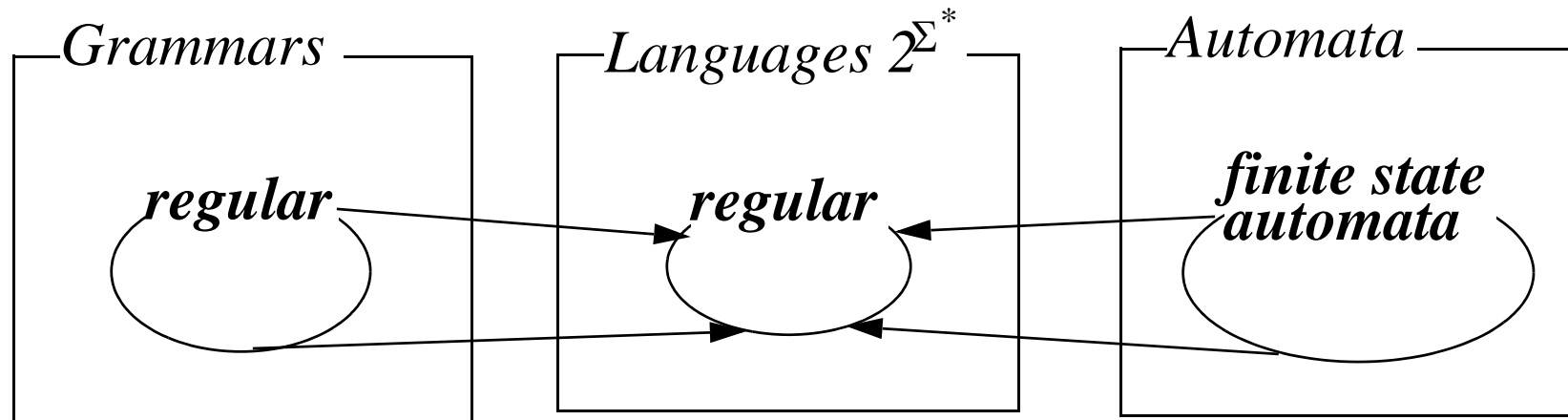
$A \rightarrow xB$ or $A \rightarrow \epsilon \in P$ where $A, B \in N$ and $x \in T^*$.

Definition A grammar $G = (N, T, P, S)$ is **regular**, if

$A \rightarrow xB$ or $A \rightarrow x \in P$ where $A, B \in N$ and $x \in T^*$.

Theorem Equivalence of finite automata and regular grammars.

proof Lem. 1 and Lem. 2.



Lem. 1 Let $A = (Q, T, \delta, q_0, F)$ be a **finite automaton (XFA)** with $\delta: Q \times \Sigma^* \rightarrow 2^Q$ and $G = (N, T, P, S)$ be a **regular grammar** where

$$(1) N = \{[q] \mid q \in Q\} \leftrightarrow Q,$$

$$(2) T = T,$$

$$(3) P = \{[q] \rightarrow x[p] \mid p \in \delta(q, x)\} \cup \{[f] \rightarrow \varepsilon \mid f \in F\}, \text{ and}$$

$$(4) S = [q_0]. \text{ Then } L(A) = L(G).$$

proof $p \in \delta^*(q, x)$, iff $[q] \Rightarrow_G^* x[p]$ is **trivial** ($? p \in \delta(q, x) \leftrightarrow [q] \rightarrow x[p]$).

If $\delta^*(q_0, x) \in F$, then $[q_0] \Rightarrow_G^* x[f], f \in F$.

$$\therefore [f] \rightarrow \varepsilon \in P.$$

$$\therefore S = [q_0] \Rightarrow_G^* x[f] \Rightarrow_G^{[f] \rightarrow \varepsilon} x.$$

$$\therefore L(A) \subseteq L(G).$$

If $S = [q_0] \Rightarrow_G^* x[p] \Rightarrow_G^{[p] \rightarrow \varepsilon} x$, then

$$p \in \delta^*(q_0, x) \text{ and } p \in F. \therefore \delta^*(q_0, x) \in F.$$

$$\therefore L(G) \subseteq L(A).$$

$$\therefore L(A) = L(G).$$

Lem. 2 Let $G_{rg} = (N, T, P, S)$ be a **regular grammar** and $A = (Q, \Sigma, \delta, q_0, F)$ be a **finite automaton** where

$$(1) Q = \{[A] \in Q \mid A \in N\} \cup \{[Ax] \in Q \mid A \rightarrow x \in P\},$$

$$(2) T = T,$$

$$(3) \delta = \{[B] \in \delta([A], x) \mid A \rightarrow xB \in P\} \\ \cup \{[Ax] \in \delta(q, x) \mid A \rightarrow x \in P\},$$

$$(4) q_0 = [S] (q_0 \leftrightarrow [S]), \text{ and}$$

$$(5) F = \{[Ax] \in Q \mid A \rightarrow x \in P\}. \text{ Then } L(G) = L(A).$$

proof $A \Rightarrow_{rg}^* xB$, iff $[B] \in \delta^*([A], x)$ is **trivial**(? 1st part of δ).

If $S \Rightarrow_{rg}^* x$, then $\exists k \geq 0$. \exists . $x = (|x|-k):x \cdot x:k$, $[Ax:k] \in \delta^*(q_0, |x|-k:x)$ and

$$\delta([Ax:k], x:k) \in F. \quad \therefore \delta^*(q_0, x) \in F. \quad \therefore L(G) \subseteq L(A).$$

If $\delta^*(q_0, x) \in F$, then $\exists k \geq 0$: $x = (|x|-k):x \cdot x:k$, $S \Rightarrow^* |x|-k:xA \Rightarrow^{A \rightarrow x:k} x$.

$$\therefore L(A) \subseteq L(G). \quad \therefore L(G) = L(A).$$

A **Rewriting system** (or **Semi-Thue system**) $R = (V^*, \rightarrow)$ where

- (1) V is a set of **configurations** (or **instantaneous descriptions; ID**), and
- (2) \rightarrow is a **finite set of relation on a free monoid V^*** ($\rightarrow \subseteq V^* \times V^*$).

A pair $(\omega_1, \omega_2) \in \rightarrow$ is called a **rule** (or **production**) of R , and denoted by $\omega_1 \rightarrow \omega_2$.

The string ω_1 is called the **left-hand side** (좌변) and ω_2 the **right-hand side** (우변) of the rule $\omega_1 \rightarrow \omega_2$.

If γ is a string in V^* that can be **decomposed** as $\alpha\omega_1\beta$ where ω_1 is **left-hand side** of the rule, then $\gamma = \alpha\omega_1\beta$ can be **written** as $\alpha\omega_2\beta$ where ω_2 is **right-hand side** of the rule, denoted as $\alpha\omega_1\beta \Rightarrow_R^{\omega_1 \rightarrow \omega_2} \alpha\omega_2\beta$.

Let $R = (V, P)$ be a rewriting system. If $r = \omega_1 \rightarrow \omega_2 \in P$. Then we define \Rightarrow_R^r (or \Rightarrow^r for short) on V^* by (\rightarrow^r is unique(**finite**) but \Rightarrow^r is **infinite**)

$$\Rightarrow_R^r = \Rightarrow_R^{\omega_1 \rightarrow \omega_2} = \{(\alpha \omega_1 \beta, \alpha \omega_2 \beta) \in V^* \times V^* \mid \alpha, \beta \in V^*\} \subseteq V^* \times V^*.$$

If $\gamma_1, \gamma_2 \in V^*$, $\gamma_1 \Rightarrow_R^r \gamma_2$, then we say that

in R γ_1 **derives** γ_2 using **rule r** and that

rule r is **applicable** to γ_1 (or **can be applied** to γ_1).

Let $\gamma \in V^*$. Then we define \Rightarrow_R^π (or \Rightarrow^r for short) **recursively** as follows

basis: $\Rightarrow_R^\varepsilon id_{V^*}$; Note that $\Rightarrow_{G'}^\varepsilon = id_{V^*}$ for any $G' = (V, P')$.

recursion: $\Rightarrow_R^\pi = \Rightarrow_R^r \Rightarrow_R^{\pi'}$ $\in P^+$ where $\pi = r\pi'$ for $r \in P$, $\pi' \in P^*$.

We can specify *initial* and *final configurations* of rewriting systems as

$$R = (V^*, \rightarrow, \mathbf{1}, \Phi)$$

(3) $\mathbf{1} \in V^*$ is an *initial configuration*, and

(4) $\Phi \subseteq V^*$ is a set of *final configurations*.

$$L(R) = \{\phi \in \Phi \mid \mathbf{1} \Rightarrow_R^* \phi \in \Phi\}.$$

A *rewriting system* for a finite automaton $A = (Q, T, \delta, q_0, F)$ is

$$R_A = (Q \times T^*, \rightarrow, (q_0, x), (F \times \Sigma^*)) \text{ where}$$

$$\rightarrow = \{(q, x) \rightarrow (p, \varepsilon) \mid p \in \delta(q, x)\}$$

$$L(R_A) = \{x \in T^* \mid (q_0, x) \Rightarrow_{R_A}^* (f, \varepsilon) \in (F \times \Sigma^*)\}.$$

A *rewriting system* for a grammar $G = (N, T, P, S)$ is

$$R_G = ((N \cup T)^*, P, S, T^*) \text{ where}$$

$$L(R_G) = \{x \in T^* \mid S \Rightarrow_{R_G}^+ x \in T^*\}.$$

Rewritings in finite automata $A = (Q, \Sigma, \delta, q_0, F)$ is

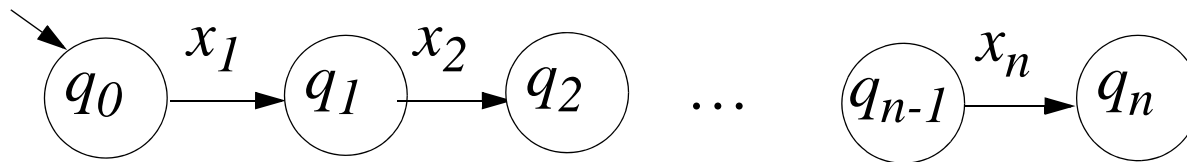
$$R_{fa} = (Q \times \Sigma^*, \rightarrow_{fa}, (q_0, x), \{(f, \varepsilon) \mid f \in F\})$$

where $(q, x) \rightarrow_{fa} (p, \varepsilon)$, if $p \in \delta(q, x)$.

Let $x = x_1x_2 \dots x_n \in \Sigma^*$ for $n \geq 0$ and $1 \leq \forall i \leq n: q_i \in \delta(q_{i-1}, x_i)$, i.e.,

$q_n \in \delta(\dots \delta(\delta(q_0, x_1), x_2), \dots), x_n)$ or

$q_1 \in \delta(q_0, x_1), q_2 \in \delta(q_1, x_2), \dots, q_n \in \delta(q_{n-1}, x_n)$, if and only if,



$$1 \leq \forall i \leq n: r_i = (q_{i-1}, x_i) \rightarrow_{fa} (q_i, \varepsilon)$$

$$(q_0, x_1x_2 \dots x_n) \Rightarrow_{fa} (q_0, x_1) \rightarrow (q_1, \varepsilon) (q_1, x_2 \dots x_n) \Rightarrow_{fa} (q_1, x_2) \rightarrow (q_2, \varepsilon) \dots$$

$$\dots \Rightarrow_{fa} (q_{n-1}, x_n) \Rightarrow_{fa} (q_{n-1}, x_n) \rightarrow (q_n, \varepsilon) (q_n, \varepsilon), q_n \in F.$$

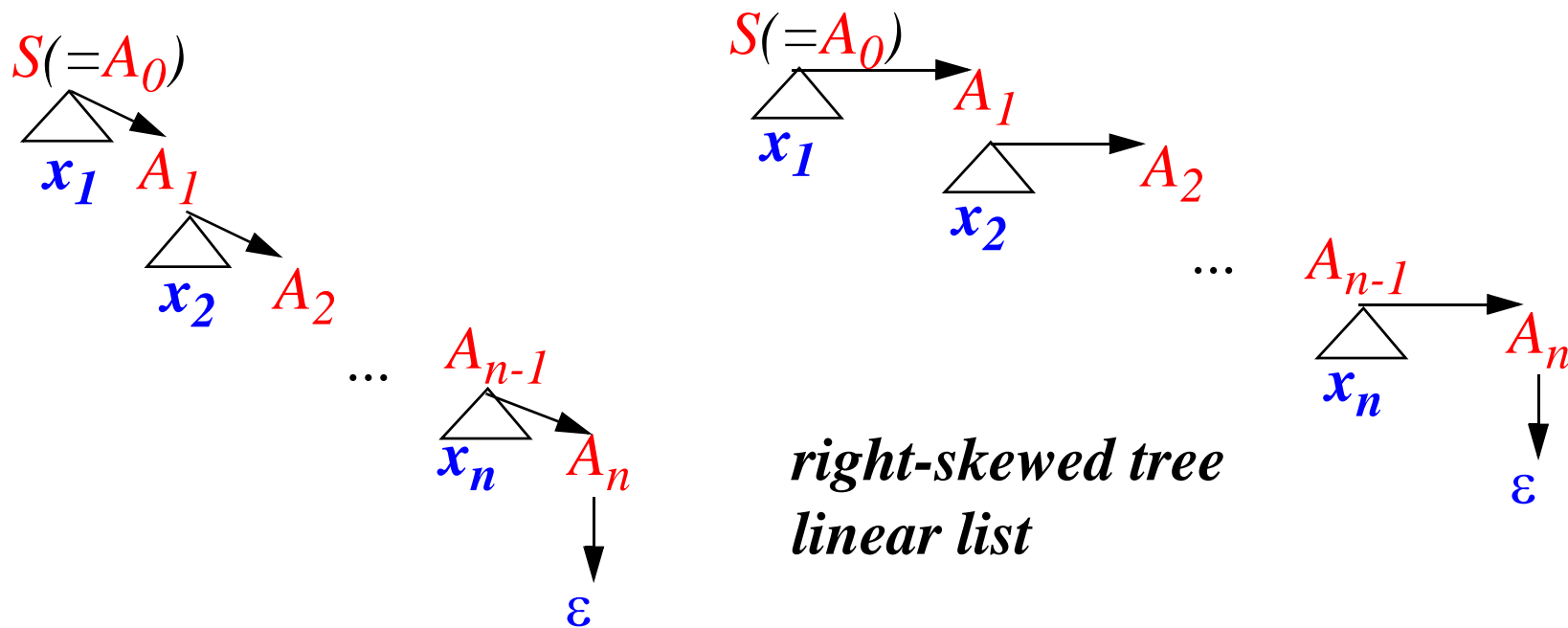
$$(q_0, x_1x_2 \dots x_n) \xrightarrow{(q_0, x_1) \rightarrow (q_1, \varepsilon)} (q_1, x_2 \dots x_n) \xrightarrow{(q_1, x_2) \rightarrow (q_2, \varepsilon)} \dots (q_{n-1}, x_n) \xrightarrow{(q_{n-1}, x_n) \rightarrow (q_n, \varepsilon)} (q_n, \varepsilon)$$

Rewritings in the regular grammar $G = (N, T, P_{rg}, S)$ is

$R_{rg} = ((N \cup T)^*, P_{rg}, S, T^*)$ $A \rightarrow_{rg} xB$, if $A \in N$, $x \in T^*$, and $B \in N \cup \{\epsilon\}$.

$1 \leq \forall i \leq n: r_i = A_{i-1} \rightarrow_{rg} x_i A_i$ and $r_{n+1} = A_n \rightarrow_{rg} \epsilon$.

$A_0 \Rightarrow_{rg} A_0 \xrightarrow{x_1} x_1 A_1 \Rightarrow_{rg} A_1 \xrightarrow{x_2} x_1 x_2 A_2 \dots \Rightarrow_{rg} A_{n-2} \xrightarrow{x_{n-1}} x_1 x_2 \dots x_{n-1} A_{n-1}$
 $\Rightarrow_{rg} A_{n-1} \xrightarrow{x_n} x_1 x_2 \dots x_n A_n \Rightarrow_{rg} A_n \xrightarrow{\epsilon} x_1 x_2 \dots x_n \in \Sigma^*$.



Compare *derivations* in (1) a regular grammar $G = (N, T, P, S)$,
 (2) *state transitions* in the **corresponding** finite automaton

$A = ([N], T, \delta, [S], F)$ where

$\delta = \{[B] \in \delta([A], x) \mid A \rightarrow xB \in P\} \cup \{f \in \delta([A], x) \mid A \rightarrow x \in P, f \in F\}$,

and (3) *rewritings* in **rewriting system**

$R_A = ([N] \times T^*, \rightarrow, ([S], xyz), F \times \{\varepsilon\})$

where $\rightarrow = \{([A], x) \rightarrow ([B], \varepsilon) \mid [B] \in \delta([A], x), [A], [B] \in [N], x \in T^*\}$.

Let $x, y, z \in T^*$, $A, B \in N$, $[A], [B] \in [N] = Q$, $f \in F$. Then

$S \quad \Rightarrow_G^* \quad xA \quad \Rightarrow_G^{A \rightarrow yB} \quad xyB \quad \Rightarrow_G^* \quad xyz \in T^*$.

grammars **generating** terminal strings

$\delta^*([S], xyz) =_A^x \delta^*([A], yz) =_A^{[B] \in \delta([A], y)} \delta^*([B], z) = f \in F$.

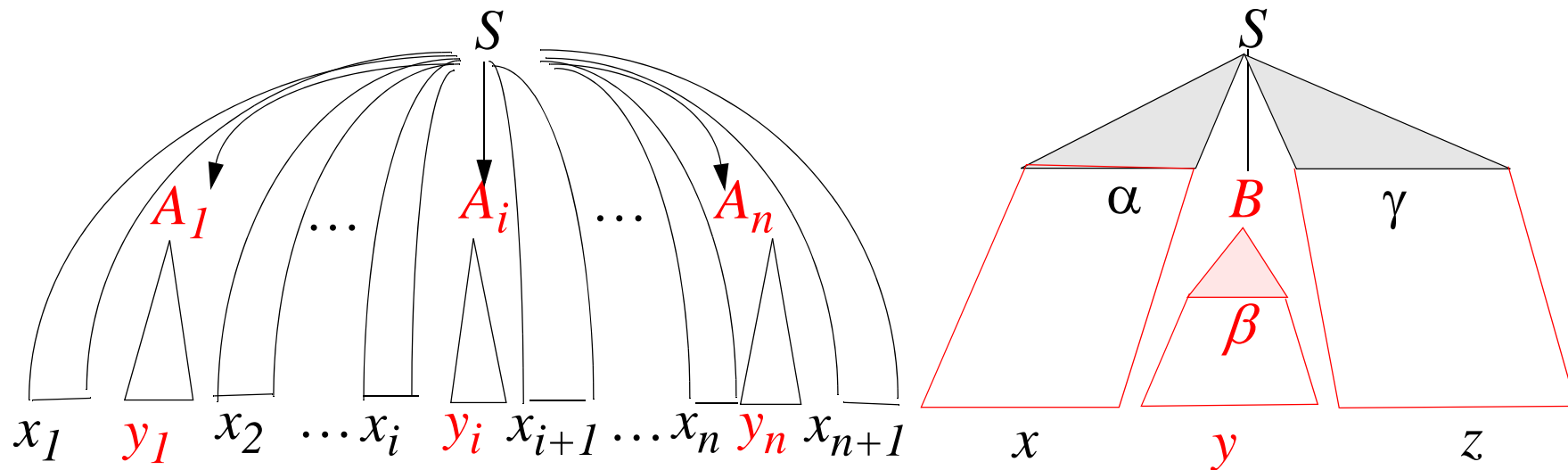
$([S], xyz) \Rightarrow_{R_A}^x ([A], yz) \Rightarrow_{R_A}^{([A], y) \rightarrow ([B], \varepsilon)} ([B], z) \Rightarrow_{R_A}^* (f, \varepsilon) \in F \times \{\varepsilon\}$.

automata **consuming** terminal strings

Let $G = (N, T, P, S)$ be a **context-free** grammar. Then

$$S \Rightarrow^* x_1 A_1 x_2 A_2 \dots x_i A_i x_{i+1} \dots x_n A_n x_{n+1} \Rightarrow^* x_1 y_1 x_2 y_2 \dots x_i y_i x_{i+1} \dots x_n y_n x_{n+1}.$$

many nonterminals in the sentential form...



Consider **two** kind of derivations \Rightarrow

leftmost derivation $\Rightarrow_{lm} \subset \Rightarrow$.

rightmost derivation $\Rightarrow_{rm} \subset \Rightarrow$.

General derivation

$$S \Rightarrow^* \alpha B \gamma \Rightarrow^{B \rightarrow \beta} \alpha \beta \gamma \Rightarrow^* \alpha y \gamma \Rightarrow^* xyz.$$

Leftmost derivation

$$S \Rightarrow_{lm}^* x B \gamma \Rightarrow_{lm}^{B \rightarrow \beta} x \beta \gamma \Rightarrow_{lm}^* x y \gamma \Rightarrow_{lm}^* xyz.$$

B is the *leftmost* nonterminal

Rightmost derivation

$$S \Rightarrow_{rm}^* \alpha B z \Rightarrow_{rm}^{B \rightarrow \beta} \alpha \beta z \Rightarrow_{rm}^* \alpha y z \Rightarrow_{rm}^* xyz.$$

B is the *rightmost* nonterminal

